

Research Impact on State Accommodation Policies for Students with Disabilities

**Martha L. Thurlow
National Center on Educational Outcomes
University of Minnesota**

April 11, 2007

Paper presented as part of a Research Symposium entitled, "Research Influences Decisions for Assessing English Language Learners and Students with Disabilities," at the 2007 annual conference of the American Educational Research Association, Chicago

Research Impact on State Accommodation Policies for Students with Disabilities

Martha L. Thurlow
National Center on Educational Outcomes
University of Minnesota

Recently, there has been more attention to the issues concerning accommodations for students with disabilities. Documentation of assessment accommodation policies for students with disabilities have also noted both an increase in the number of states with policies for students with disabilities and an increase in the sophistication of the policies. Research findings for accommodations provided to students with disabilities have been complex, yet still have pushed states to make a better determination of their policies and practices. This paper summarizes research findings on accommodations for students with disabilities and discusses the nature of accommodation policies, including recent specifications about the scoring and aggregation of test scores and the use of “access assistants.”

There are many definitions of testing “accommodations,” yet all now converge on the same concepts. In general, any type of change to testing materials, setting, or procedures that does not alter what is being measured is considered to be an accommodation. Some researchers refer to the elimination of construct irrelevant variance, in other words, the variance associated with extraneous features of test administration (Fuchs, Fuchs, Eaton, Hamlett, & Karns, 2000). Demonstrations of test changes that maintain the validity of the construct that an assessment is attempting to measure have been challenging for the field, and numerous approaches to research have been pursued (Thurlow, McGrew, Tindal, Thompson, Ysseldyke, Elliott, 2000; Sireci, Scarpati, & Li, 2005; Tindal, 1998), including single subject designs, “boost” studies, and “differential boost” studies.

Accommodations continue to be a topic of heightened interest for a number of reasons. First and foremost, it is no longer a question as to whether students with disabilities need to be participants in national, state, and district assessments. Many of those students with disabilities who do participate require accommodations; data from the 2003-2004 Annual Performance Reports submitted by states to the federal government indicated that an average of 61-66% of school-age students with disabilities used accommodations during assessments (Thurlow, Moen, & Altman, 2006). Participation in national, state, and district assessments is required by both the federal special education law, the Individuals with Disabilities Education Act (IDEA) of 1997 and as reauthorized in 2004, and by the federal Elementary and Secondary Education Act, now known as the No Child Left Behind Act (NCLB), which requires participation in state assessments and the National Assessment of Educational Progress (NAEP).

The challenges surrounding accommodations, of course, are determining which changes result in alterations to what is being measured. Historically, this concern has not always been a top priority. For some time, the major focus was on the participation of students in the assessments – “accommodations” were seen as an avenue to participation, and the distinction between whether a change compromised the construct being assessed was not necessarily the major factor of interest (Thurlow, 2007). It has taken time for the field to adjust its focus to clearly distinguish between “accommodations” and “modifications” (those changes in materials, setting, or procedures that result in changes to the construct measured), and for research and policy to focus as well on these concepts.

As the field has moved forward, both in terms of policy development and in terms of research, it has been pushed at the same time by federal activities in which state assessment systems are undergoing reviews by their peers. One aspect of the peer reviews is that students with disabilities will participate

with accommodations, and thus, these receive review as part of the assessment system. Specific aspects of the peer review guidance (U.S. Department of Education, 2004) that address accommodations for students with disabilities are:

- 4.3 Has the State ensured that its assessment system is fair and accessible to all students, including students with disabilities and students with limited English proficiency, with respect to each of the following issues:
- (a) Has the State ensured that the assessments provide an appropriate variety of accommodations for students with disabilities?
 - (d) Does the use of accommodations . . . yield meaningful scores?
(p. 37)
- 4.6 Has the State evaluated its use of accommodations?
- (a) How has the State ensured that appropriate accommodations are available to students with disabilities and that these accommodations are used in a manner that is consistent with instructional approaches for each student, as determined by a student's IEP or 504 plan?
 - (b) How has the State determined that scores for students with disabilities that are based on accommodated administration conditions will allow for valid inferences about these students' knowledge and skills and can be combined meaningfully with scores from non-accommodated administration conditions?
(p. 40)

The federal requirements have ratcheted up the need for states to attend to the research and to ensure that their students are using accommodations during assessments that are producing valid scores. In fact, during the proposal for new regulations in 2005, the Department of Education attempted to confirm in regulations practice that had been imposed through non-regulatory guidance – that being that any student who participated in assessments in a way that produced invalid test results (and this included using changes in testing procedures considered by the state to be modifications) would no longer be considered participants in the assessment. For *No Child Left Behind* Adequate Yearly Progress (AYP) accountability purposes, which require that each subgroup have at least 95% participation for the school or district to be eligible for meeting AYP, this posed a serious threat. The specific words in the Notice of Proposed Rule-Making (December 15, 2005) were as follows:

Tests administered with accommodations that do not maintain test validity are not measuring academic achievement under the State's assessment system. Under the reauthorized IDEA, each IEP now must indicate "appropriate accommodations that are necessary to measure the academic achievement and functional performance of the child on State and district wide assessments." State and LEA guidelines thus need to identify, for IEP teams, those accommodations that will maintain test validity. Similarly, under Title I, the concept of "appropriate accommodations" in the context of assessments must be thought of as accommodations that are needed by the individual child and that maintain test validity. The Title I regulations would only consider a student to be a participant for AYP purposes if his or her assessment results in a valid score.

These developments suggest that there is a whole new set of pressures in place on research and on policy. They are in the works at the same time that research and policy are continuing to react to the need for more information to respond to what is a critically complex area of study and policy making.

The purpose of my paper is to stand back again and look at what we know at this point. First, what does the research evidence tell us – both what has been happening in the research on test changes for students with disabilities and what can we glean from that research. Second, what are state

policies like, and how have they changed with respect to some specific accommodations? Finally, I want to focus more specifically on just a couple aspects of state policies – how they are dealing with the aggregation of scores in some specific cases, and how they are dealing with some specific kinds of accommodations – ones that involve humans – what we at NCEO have called access assistants.

Research Findings

Several years ago we were begging for research funds to support research on accommodations. The U.S. Department of Education responded to those calls and funded research for several years; in addition, the National Center on Education Statistics (NCES), and some state departments of education have conducted their own studies of interest.

Published research on accommodations has been summarized by NCEO, in follow-up to initial work by Tindal and Fuchs (1999) to show the nature of the research that has been conducted across time (Thompson, Blount, & Thurlow, 2002; Johnstone, Altman, Thurlow, & Thompson, 2006; Zenisky & Sireci, in press). For example, there has clearly been an increase in research over the years, with the average per year increasing from about 4 studies of test changes in the early 1990s to about 16 per year in the mid-2000s:

Years	Number of Studies	Average Per Year
1990 through 1992	11	3.7
1993 through 1995	18	6.0
1996 through 1998	29	9.7
1999 through 2001	46	15.3
2002 through 2004	49	16.3
2005 through 2006	32	16.0

The research that has been conducted since 2002 is primarily directed at determining the effects of the use of accommodations on test scores of students with disabilities and at investigating the effects of accommodations on test score validity. The research methods are more often experimental or quasi-experimental than other techniques such as surveys, interviews, or case studies (see Table 2).

Method	Number of Studies	
	2002-2004*	2005-2006**
Experimental or Quasi-experimental	21 (43%)	18 (56%)
Review of extant data	17 (35%)	6 (19%)
Survey/Interview	7 (14%)	5 (16%)
Meta-analysis	2 (4%)	0 (0%)
IEP intervention	1 (2%)	0 (0%)
Product evaluation	1 (2%)	0 (0%)
Other (i.e., case studies, observation)	0 (0%)	3 (9%)

* From Johnstone, Altman, Thurlow, & Thompson (2006).

** From Zenisky & Sireci (in press).

Most Commonly Researched Accommodations. The specific accommodations that have been the focus of research on accommodations during the past five years are show in Table 3. Oral administration and extended time are the most frequently researched accommodations. This was also true in the 1999-2001 summary of accommodations research; other types of accommodations were included then that are not included in the past five years for students with disabilities (e.g., simplified

language, dictated response, frequent breaks) as well as other differences (e.g., more studies on small group and multiple accommodations) (Thompson, Blount, & Thurlow, 2002).

Type of Accommodation		Number of Studies	
		2002-2004*	2005-2006**
Presentation	Oral Administration	11	5
	Computer Administration	5	3
	Video	2	0
	Scrolling vs. Paging	0	1
	Large Print	1	0
	Dictionary Use	1	0
	Braille	1	0
Timing/Scheduling	Extended Time	7	7
	Multiple Day	1	1
	Separately Timed Sessions	0	1
Response	Dictated Response	1	0
	Calculator	1	0
Setting	Small Group	0	1
Technological Aid		2	0
Multiple Accommodations		11	0
N / A (Other)		5	19

* From Johnstone, Altman, Thurlow, & Thompson (2006). Percentages based on

** From Zenisky & Sireci (in press). Percentages based on 21 studies.

The major findings related to the most commonly researched accommodations are interesting, and point out some of the challenges of conducting research on the effects of accommodations, and of drawing conclusions from that research. Table 4 reproduces the short summary of findings that the literature reviewers listed for each of the studies on oral administration (see Johnstone et al., 2002 and Zenisky & Sireci, in press).

Author	Results
Barton (2002)	The results indicate that a similar construct was measured among students with and without disabilities taking the regular form. The results also indicate that a similar construct was measured among students with and without disabilities taking the oral accommodation form.
Barton & Huynh (2003)	The study indicates that when errors are used as an extra factor in exploring the nature of proficiency, the reading construct varies only slightly across disability groups. The results indicate that it is safe to apply the same meaning to test scores for these groups even when the test is administered under different accommodations.
Bolt & Bielinski, (2002)	The read aloud accommodation did not appear to improve score comparability for students with reading disabilities when compared to students without disabilities. More items displaying differential item functioning (DIF) were identified for those who received the accommodation than for those who did not receive the accommodation.
Helwig, Rozek-Tedesco, & Tindal (2002)	Elementary students with disabilities tended to perform better under the read aloud condition; elementary general education students did not appear to receive a similar benefit from the accommodation. For middle school students, no significant interactions were found.
Huynh, Meyer, & Gallant (2004)	It was found that the test structure remained rather stable across the three groups. Controlling for student background variables, disabled students under oral administration performed better than disabled students on the non-accommodated format. On the non-accommodated format, students with disabilities fared worse than general education students.
Janson (2002)	Students who received special education services and received accommodations experienced significant gains in scores in science and social studies in the year they were initially granted accommodations. There were substantial gains in science and social studies in 2000 for students

	initially receiving accommodations. There were significant gains in social studies and math scores in 2001 for students initially receiving accommodations.
McKevitt & Elliott (2003)	The use of the read-aloud accommodation did not significantly improve the test performance of either group of students. Teachers as a group had neutral attitudes about testing and testing accommodations.
Meloy, Deville, & Frisbie (2002)	Analyses revealed that students in both groups (LD-R and non-LD) achieved significantly higher test scores with the read aloud test administration.
Tindal (2002)	A main effect for both student classification and test administration was found for the elementary school students: Low achieving students outperformed students with IEPs and both groups benefited with a video-taped administration. For middle school students, a main effect was found for student classification; however, no main effects for the type of test administration (video versus standard).
Weston (2003)	The findings revealed a statistical difference between the tests, and also between the two groups of students. Students with learning disabilities who are poor readers gained the most from the read-aloud accommodation. Results also suggest that the results on the accommodated test better match the teacher's estimations of the student's mathematical abilities.
Woods (2004)	The investigation found a low level of candidate need for a reader with candidate reading age and self-prediction being unreliable indicators of this need.
Dolan et al. (2005)	Scores on the computerized-oral test were significantly increased over paper scores when passages were longer than 100 words in length.
Fletcher et al. (2006)	Only students with disabilities benefited from the accommodations, showing a significant increase in average performance and a 7-fold increase in the odds of passing; results supported the interaction hypothesis.
Huynh & Barton (2006)	After controlling for major background variables, the performance of students with disabilities under oral administration conditions was comparable to those of students without disabilities who took the test under regular administration conditions. The internal structure of the HSEE test remained stable across students with and without disabilities.
Schnirman (2006)	No statistically significant differences were found between performance of students with and without disabilities.

* Quoted from Johnstone, Altman, Thurlow, & Thompson (2006) and Zenisky & Sireci (in press).

Clearly the results are complicated by the inclusion of different groups of students, the study of different content areas, the use of different media for presenting the accommodation (person vs video vs audio tape), and by other refinements (such as the length of the passage to be read). Nevertheless, research on the oral administration of tests continues as an area of high interest, particularly for reading assessments where students with disabilities that impact reading are struggling to show their knowledge and skills on other standards that are measured by these assessments (see National Accessible Reading Assessment Projects, 2006). In their review of studies prior to 2005-06, Sireci et al. (2005) had concluded that for the oral administration "half the studies that focused on oral accommodations on math tests found positive effects. However, studies that focused on other subject areas found either no gains for SWD or similar gains for SWD and for students without disabilities" (p. 484). The results overall were considered "promising," but unclear.

Research on extended time, the other test change that has received considerable attention through research, is somewhat different. Table 5 reproduces the short summary of findings that the literature reviewers listed for each of the studies during the past five years that focused on oral administration (see Johnstone et al., 2002 and Zenisky & Sireci, in press).

Author	Results
Bridgeman, Cline, & Hessinger (2004)	Extra time added about 7 points to verbal scores and 7 points to quantitative scores. The accommodation appeared to have a greater impact on the quantitative scores of lower ability examinees.
Buehler (2002)	Results indicated that students with learning disabilities did not use significantly more time on the CAT/5, even when given the option. Students with disabilities did not receive any differential benefit from the use of the additional time accommodation. Although there were no differences in the reliability of the CAT/5 due to the accommodation of additional time, the validity of the CAT/5 was lower for students with learning disabilities

	who received additional time. The CTOPP was not found to be a useful predictor of students that would benefit from additional time on the CAT/5.
Cahalan, Mandinach, & Camara (2002)	In general, the revised SAT was to be positively correlated with FGPA for students who took the test with extended time accommodations for a learning disability. SAT scores were fairly accurate predictors of FGPA for students with learning disabilities. In the majority of cases when HSGPA was used along with SAT test scores, the predictive validity of FGPA was increased.
Dempsey (2003)	The verbal comprehension index was identified as the score that most closely predicts LSAT performance. This study found that scores earned under accommodated conditions are better than those earned under standard conditions.
Elliott (2003)	The performance of students with disabilities was highly similar to the performance of students without disabilities under standard time and extended time testing conditions. Overall, the provision of the accommodation, extended time, did not significantly improve scores of students with disabilities on the math test.
Elliott (2004)	The scores achieved in the extended time condition were higher than the scores achieved in the standard condition for all groups. However, the scores of students with disabilities did not improve significantly more than those of the students without disabilities when given extra time. A large proportion of survey respondents across all three groups expressed approval of the extended time condition.
Thornton, Reese, Pashley, & Calessandro (2002)	Overall, results suggest that LSAT scores earned under the nonstandard time condition are not comparable to LSAT scores earned under standard timing conditions. Results for individual subgroups were consistent with the overall group result.
Antalek (2005)	The majority of the subjects took additional time and their scores on the task improved significantly, indicating a relationship between learning disabilities and the completion of academic tasks within an allotted time frame.
Baker (2006)	The group that used extended time accommodations had an average first year GPA that was 0.39 points higher (statistically significant) than the group that did not use accommodations. The use of extended time accounts for 11% of variance in full year GPA and 7% of overall GPA.
Bolt & Ysseldyke (2006)	Read-aloud accommodations and extended time were found to be associated with a comparable level of DIF relative to the use of read-aloud only, and these results were consistent across both reading and math.
Cohen, Gregg, & Deng (2005)	Some items exhibited DIF under accommodated conditions, but students for whom items functioned differently were not accurately characterized by their accommodation status but rather content knowledge.
Lesaux, Pearson, & Segel (2006)	Under timed conditions there were significant differences between performance of students with disabilities and students without disabilities; all of the students with disabilities benefited from extra time, but students without disabilities performed comparably under timed/untimed conditions. Also, students with disabilities (less severe) performed comparably to students without disabilities in untimed conditions.
Mandinach, Bridgeman, Cahalan-Laitusis, & Trapani (2005)	Results indicated that time and a half with separately timed sections benefits students with disabilities and students without disabilities, though some extra time improves performance and too much may be detrimental. Extended time benefits medium and high ability students but provides little or no advantage to low-ability students.

* Quoted from Johnstone, Altman, Thurlow, & Thompson (2006) and Zenisky & Sireci (in press).

The studies on extended time show a somewhat more consistent picture. In their recent review of research on extended time (based on the most recent studies), Zenisky and Sireci (in press) concluded:

The results for extended time, as the most frequently researched accommodation in the 32 studies considered here, are generally consistent with the previous literature, where extended time has been shown to have a positive effect on the scores of students with disabilities. However, the emerging trend in elementary and secondary education toward the use of untimed tests for all students (as part of a larger strategy of integrating universal test design noted by Sireci, et al. (2005)), if it continues, may yet minimize the need for further study of the benefits of extended time test accommodations. (p. 15)

This statement confirms the observation that research and policy are moving along somewhat parallel tracks, possibly influencing each other, in ways that will increase the likelihood that test results for students who receive accommodations are valid.

Most Commonly Allowed Accommodations. The accommodations that have been studied in the past five years are not necessarily the most commonly allowed accommodations (especially if extended time is no longer considered an accommodation but is standard practice). Bolt and Thurlow (2004) reported on the research on the most commonly allowed accommodations in 1999 (Braille, dictated response, large print, extended time, sign language interpreter), and found that in the 36 identified studies on these, there was mixed support and nonsupport for the accommodations for students with disabilities. When Bolt and Thurlow selected accommodations to include in their study, they included accommodations both with and without limitations.

In terms of continuing to examine research findings, it might be useful to make a distinction between those test changes that are allowed by states without restrictions and those test changes that are allowed with restrictions. The test changes that, according to state policies in 2003 and 2005, were the most often allowed without restrictions, are shown in Table 6. In this table, it is also indicated whether some states allowed the test change **with** restrictions and whether some states prohibited the test change. For some test changes, more than 5 states (10% of the 50 states) altered their policies – so that the number of states in 2005 was different from the number in 2003. This occurred for **Braille Edition** (which increased in the number of states allowing without restrictions, and decreased in the number of states allowing with restrictions), **separate room** (which decreased in the number of states allowing without restrictions and also was mentioned by fewer states), and **time beneficial to student** (which increased in the number of states allowing without restrictions and also with restrictions – showing an increase in the total number of states mentioning the test change). Of interest is the fact that few studies during this time frame examined the accommodations where dramatic changes were made in policy. This is true even if one goes back to the research before the time period of the current policies – if one assumes that there is more of a lag between research and policy.

Table 6. Test Changes Most Often Allowed Without Restrictions*

Test Change	States Allowing Without Restrictions		States Allowing With Restrictions		States Prohibiting	
	2003	2005	2003	2005	2003	2005
Large Print	47	48	2	0	0	0
Individualized Administration	46	45	0	0	0	0
Small Group Administration	47	45	0	0	0	0
Magnification	41	42	0	0	0	0
Braille Edition	38	46	11	2	0	0
Separate Room	38	31	0	1	0	0
Write in Test Booklet	35	35	4	5	0	0
Time Beneficial to Student	35	41	0	4	0	0

*The 2003 information is from Clapper, Morse, Lazarus, Thompson, & Thurlow (2005). The 2005 information is from Lazarus, Thurlow, Lail, Eisenbraun, & Kato (2006).

Another way to look at research on accommodations in the past five years is in terms of whether it has addressed those accommodations that are most frequently **allowed with restrictions**. These test changes tend to be the accommodations that are more controversial, and that need specifications placed on them (e.g., states allow them to be used in one situation but not another; states allow them to be used by some students but not other students). Table 7 shows the accommodations that are most often allowed with restrictions (also including an indication of the numbers of states that allow without restrictions and the numbers of states that prohibit). These are the accommodations that have received attention, either currently or in the past, and many of the changes reflect both a recognition of research findings and a policy push.

Table 7. Test Changes Most Often Allowed Without Restrictions *

Test Change	States Allowing Without Restrictions		States Allowing With Restrictions		States Prohibiting	
	2003	2005	2003	2005	2003	2005
Oral Administration/Read Aloud	3	8	44	37	0	0
Calculator	15	19	28	22	1	0
Proctor/Scribe	32	37	17	11	0	0
Extended Time	29	41	16	4	2	0
Sign Interpret Questions	13	8	29	25	0	0

Note: The 2003 information is from Clapper, Morse, Lazarus, Thompson, & Thurlow (2005). The 2005 information is from Lazarus, Thurlow, Lail, Eisenbraun, & Kato (2006).

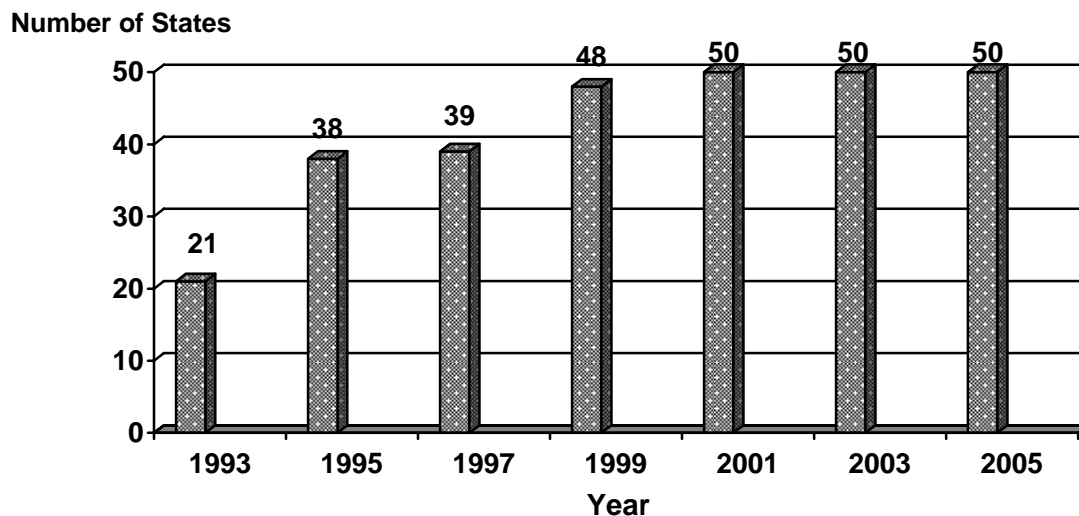
Although it might be desirable to link increases or decreases in certain allowed with or without restrictions of test changes directly to the research findings, it is more likely that we are dealing with a unique set of merging factors pushing for policy changes. It is important to look in general at accommodation policies and what has been happening with respect to their evolution and changes in them to get a sense of the effects of research on the changes in policies related to test changes.

In contrast to most research, which has had general effects, some research has had specific policy impact. For example in large states with enough students with specific disabilities research has been undertaken to address specific research questions. For example, Fletcher et al. (2006) examined accommodations specifically designed to minimize the effect of word recognition difficulties for students with word decoding problems and found that only the students with disabilities benefited. In other words, the interaction hypothesis was confirmed. Their study led to specific changes in the Texas accommodation policies. Not every state is large enough to conduct this type of studies, or has the structures and resources needed to do so.

Accommodation Policy Shifts

There have been a number of shifts in accommodation policies over time. These include the steady but dramatic increase in the number of states with accommodation policies/guidance documents, from 21 in 1993 to 50 for the past several years (see Figure 1).

Figure 1. Number of States with Accommodation Policies Over Time



Nature of States' Accommodation Policies. States' accommodation policies themselves have changed in several ways also. When NCEO first started studying states' policies, we had to contact states to obtain copies of their documents, and even in 1995, we were still able to reproduce all of the accommodation policies and guidance in one report that was less than 175 pages long. This report quoted all the relevant parts of the policies in all of the states that had them (Thurlow, Scott, & Ysseldyke, 1995). Today, and for the past several years, states' accommodation policies have blossomed. They are now available on the state's Web site in nearly every state, and each one in several states is more than 175 pages long.

Sophistication of Policies. State accommodation policies are much more complex than ever before. When NCEO first started summarizing accommodation policies, we simply indicated an X for a test change that was "allowed," a blank for one that was not mentioned by the state, and a P for a prohibited test change. Each time or couple times we summarized policies an adjustment was made to better reflect the increasing complexity of the policies themselves. Table 8 reflects the coding changes that have occurred over time, and in turn provides a glimpse of the complexity and sophistication of the policies.

Year	Coding System
1993	X = allowed; O = prohibited
1995	Lists of states with certain accommodations
1997	X = allowed; O = prohibited
1999	X = allowed; O = prohibited
2001	X = allowed; X* = score not aggregated if used; O = prohibited
2003	A = allowed; AC = allowed in certain circumstances; AI = allowed with implications for scoring and/or aggregation; P = prohibited
2005, 2007	A = allowed; A* = allowed but called nonstandard (with no implications for scoring or aggregation); AC = allowed in certain circumstances; AI = allowed with implications for scoring and/or aggregation; P = prohibited

Implications for Aggregation of Scores. The type of sophistication reflected in recent years indicates that states are attending to what happens to scores and the aggregation of scores (including the reporting of scores) when test changes have been introduced. Clarity about the effects of the test changes on the validity of test results clearly is of concern to states. This does not mean that all states are in agreement with respect to aggregation for many accommodations.

NCEO introduced the code AI = allowed with implications for scoring and/or aggregation in 2003. This was a modification of the code used in 2001, which indicated a more rigid interpretation (score not aggregated if used). In fact, what is frequently observed is that the implications for scoring or aggregation may depend on specific circumstances, such as the content of the assessment or the assessment itself. Table 9 shows several "allowed with implications for scoring and/or aggregation" circumstances from 2005 – for the proctor/scribe accommodation – to give a sense of what the specifications are like.

State	Specifications
Arizona	Dictating to a scribe is considered a non-standard accommodation when used on the writing portion of Arizona's instrument to Measure Standards (AIMS).
California	Proctor/Scribe – allowed with implications for scoring if used on the CST, CAHSEE, or CELDT.
Hawaii	Proctor/Scribe – Must be in an individual setting; Allowed with implications for scoring if used on any test.
Massachusetts	Proctor/Scribe – Considered non-standard if used on the ELA Composition Test (may alter what the test measures).

Oregon	Proctor/Scribe – Considered a modification if used on writing test (not considered part of standard administration; scores obtained under modified conditions do not allow students to meet content and achievement standards and the scores will not appear in school and district group statistics).
Utah	Proctor/Scribe – Considered a modification on all tests except for the Iowa tests.
Vermont	Proctor/Scribe – Allowed with implications for scoring if used on the writing test.

Thus, the notion of states' policies moving forward with research is a practical reality. As indicated in Table 9, some states appear to be more tuned into variations in content than others. Previous research had not effectively isolated the effects of scribes (Elliott, Bielinski, Thurlow, DeVito, & Hedlund, 1999; Fuchs et al., 2000; Schulte, Elliott, & Kratochwill, 2001), but had suggested that some students who had used scribes had obtained improbably high scores (Koretz & Hamilton, 2000), a finding that had affected policy in the past. Moving forward, some states attempted to address policies through other means – controlling sources of variation in the administration of the accommodations themselves.

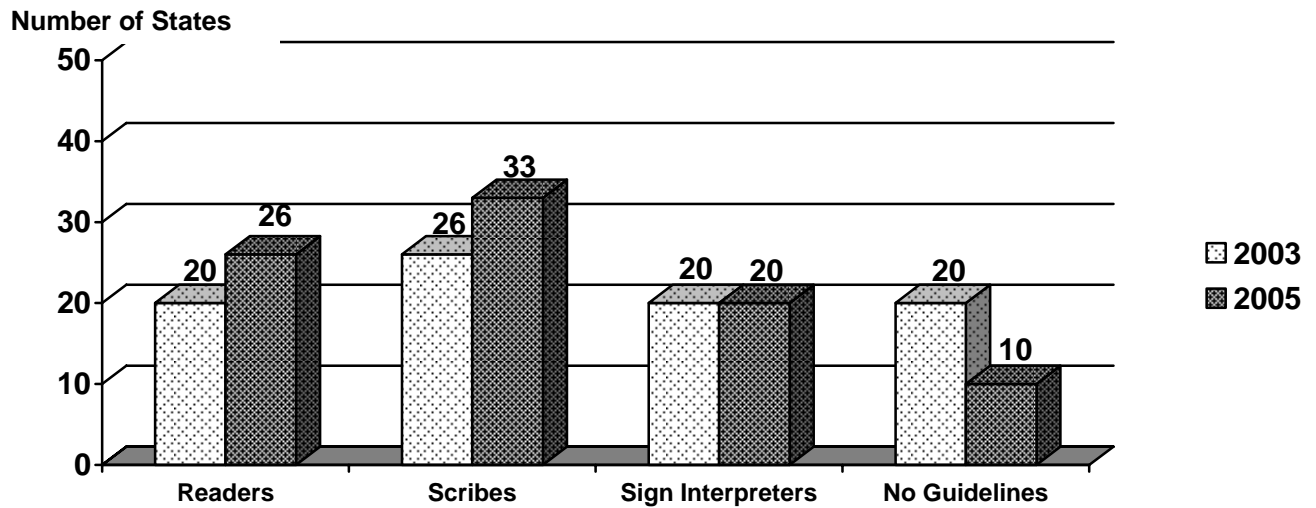
Access Assistants. Many states have attempted to increase the validity of specific accommodations that in the past have been a challenge. Most often these are the ones that had been designated as “allowed with implications for scoring and/or aggregation.” In many cases, these accommodations involve a human as an element of the accommodation – such as a human reader for the student (reader), a human who signs the test questions to the student (sign language interpreter), or a human who writes for the student (scribe). Occasionally states attempt to move to machines to take over the functions of the humans (such as in allowing text-to-speech assistive technology). Regardless of these attempts, the need often remains for some students to have a human involved in their accommodations – NCEO has called these individuals “access assistants” in recent attempts to study the implementation and validity of these accommodations. The term “access assistant” encompasses readers, scribes, sign language interpreters, and other accommodations that involve a human as part of the accommodation itself.

The introduction of a person who interacts with the student raises a number of validity issues. There is always the concern that the human will inadvertently (or not so inadvertently) give cues to the student about the correct answers. It has become increasingly important for states to attend to the validity of results from accommodations that involve access assistants. One avenue for doing this is to require that these individuals meet specific qualifications, that they conduct themselves in specific ways, and that they are open to scrutiny.

Two studies have now been conducted to look at the guidelines that states have for readers, scribes, and sign language interpreters – one in spring 2003 (Clapper et al., 2005), and one in 2005 (Lazarus et al., 2006). The written guidelines that were sought were ones that defined the role, prescribed conditions or qualifications, or provided specific guidance for implementation. Quality criteria were not imposed for counting guidelines, nor were any guidelines evaluated.

Figure 2 shows the numbers of written guidelines for access assistants that were found for the 50 states in 2003 and 2005. As is evident in this figure, the number of guidelines increased for both readers and scribes, but stayed the same for sign language interpreters. Overall, the number of states without guidelines for any of the three types of access assistants (see last column in Figure 1) decreased from 20 in 2003 to only 10 in 2005). This figure reveals that while states increasingly are adopting guidelines for their access assistants, they are not necessarily doing it for all of them – they are selecting those that they are most concerned about, or perhaps those that they believe that they are able to control. This clearly is most often the scribe.

Figure 2. States with Written Guidelines for Access Assistants in 2003 and 2005



Despite the focus that states have placed on scribes in their guidance, there appears to still be a way to go – not just in the numbers, but in the content. While NCEO’s analysis did not attempt to conduct any qualitative analysis of the content of the guidelines, it is important to look at them with somewhat of an evaluative eye. A more indepth analysis of the information included in states’ guidelines for access assistants reveals that not all guidelines give information on:

- Specific procedures for providing the accommodation to the student
- Preparation for providing the accommodation (including studying the assessment – when and how to do)
- Requirements for qualifications and characteristics of the access assistant

Looking just at information for the scribe, Clapper, Morse, Thompson, & Thurlow (2005) found that 20 states included information that addressed the specific procedures to follow for providing the accommodation to the student, including what to do to prepare for providing the assessment. Only 13 states provided specific requirements for the qualification or characteristics of the access assistants.

With NCEO’s 2007 update of states’ participation and accommodation policies, there will again be an examination of the status of states’ access assistance guidelines. These are a critical piece in the validity of results from accommodated assessments, and it is likely that states are more attuned than ever to the results from assessments administered with access assistants.

Discussion

As policy and practice continue to push forward the inclusion of students with disabilities in large-scale assessments, with a significant number of these students using accommodations to participate in the assessments and to show what they know and are able to do, the field finds itself still plagued by concerns about whether results from accommodated assessments produce valid scores. In many ways, the field has been learning by leaps and bounds as it has refined its conceptual frame and its methodologies for examining questions of the validity of test results when there are changes in the test administration. Still, many questions remain, and with the diversity of students with disabilities,

and the complexity of accommodations, one begins to realize that solutions will have to come from a marriage of research, rationales, and wisdom.

Although we have determined how to do the best research, we have to recognize that the research is thwarted by at least two real problems. First, students with disabilities are not simple beings, and they do not need just one or two accommodations. We have not yet addressed how to handle the interaction of the accommodations and what these interactions might mean for the results that have been obtained – or for the research that is underway. Second, we have to address what is happening in the application and the implementation. There is a whole new wave of research that is addressing accommodations implementation and the effects of accommodation as they are implemented. We must attend to this research as it moves forward, to the assumptions that it brings to the field, and to the directions it can take us. It will continue to be essential to attend to training and IEP decision making as well because they are an essential part of policy implementation and of research. These are not neat and clean topics, but they affect our research.

All the research that has been conducted on the effects of accommodations provides bits and pieces to a bigger puzzle, which we are slowly putting together. As we do this, we are also moving along in the direction of identifying things that can be done to assessments to make them better from the beginning. This has been the concept of universal design. We cannot let ourselves become distracted by thinking that by inserting accommodations within computer-based testing that we have achieved universal design – the concept is something much bigger and grander. It is about starting from the very beginning and thinking about students with disabilities and students who are English learners as we design our tests. Research and policy on accommodations have shown us how important it is that we pursue that notion – that we determine ways to make our assessments more accessible and that we pursue the notion of universal design, so that fewer students need accommodations to be able to participate in assessments that produce valid measures of what they know and can do.

References

- Antalek, E. E. (2005). The relationships between specific learning disability attributes and written language: A study of the performance of learning disabled high school subjects completing the TOWL-3. Dissertation Abstracts International, 65 (11), 4098 A. Retrieved August 5, 2006 from Digital Dissertations database.
- Baker, J. S. (2006). Effect of extended time testing accommodations on grade point averages of college students with learning disabilities. Dissertation Abstracts International, 67 (1), 574 B. Retrieved August 5, 2006 from Digital Dissertations database.
- Barton, K.E. (2002). Stability of constructs across groups of students with different disabilities on a reading assessment under standard and accommodated administrations (Doctoral dissertation, University of South Carolina, 2001). Dissertation Abstracts International, 62/12. 4136.
- Barton, K.E., & Huynh, H. (2003). Patterns of errors made by students with disabilities on a reading test with oral reading administration. Educational and Psychological Measurement, 63 (4), 602-614.
- Bolt, S., & Bielinski, J. (2002). *The effects of the read aloud accommodation on math test items*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Bolt, S.E., & Thurlow, M.L. (2004). A synthesis of research on five of the most frequently allowed testing accommodations in state policy. Remedial and Special Education, 25 (3), 141-152.
- Bolt, S. E., & Ysseldyke, J. E. (2006). Comparing DIF across math and reading/language arts tests for students receiving a read-aloud accommodation. Applied Measurement in Education, 19 (4), 329-355.
-

- Bridgeman, B., Cline, F., & Hessinger, J. (2004). Effect of extra time on verbal and quantitative GRE scores. *Applied Measurement in Education, 16* (3), 191-205.
- Buehler, K.L. (2002). Standardized group achievement tests and the accommodation of additional time (Doctoral dissertation, Indiana State University, 2001). *Dissertation Abstracts International, 63/04*, 1312.
- Cahalan, C. Mandinach, E., & Camara, W.J. (2002). *Predictive validity of SAT I: Reasoning test for test-takers with learning disabilities and extended time accommodations*. New York: The College Reporting Board.
- Clapper, A.T., Morse, A.B., Lazarus, S.S., Thompson, S.J., & Thurlow, M.L. 2003 *state policies on assessment participation and accommodations for students with disabilities* Synthesis Report 56). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Cohen, A. S., Gregg, N., & Deng, M. (2005). The role of extended time and item content on a high-stakes mathematics test. *Learning Disabilities Research & Practice, 20* (4), 225-233.
- Dempsey, K.M. (2004). The impact of additional time on LSAT scores: Does time really matter? The efficacy of making decisions on a case-by-case basis Doctoral dissertation, La Salle University, 2004). *Dissertation Abstracts International, 64/10*, 5212.
- Elliott, J.L., Bielinski, J., Thurlow, M.L., DeVlto, P., & Hedlund, E. (1999). *Accommodations and the performance of all students on Rhode Island's performance assessment* (Rhode Island Assessment Report 1). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Elliott, S.N., & Marquart, A.M. (2003). *Extended time as an accommodation on a standardized mathematics test: An investigation of its effects on scores and perceived consequences for students with varying mathematical skills*. Madison, WI: University of Wisconsin-Madison, Wisconsin Center for Education Research.
- Elliott, S.N., & Marquart, A.M. (2003). Extended time as a testing accommodation: Its effects and perceived consequences. *Exceptional Children, 70*(3), 349-367.
- Fletcher, J.M., Francis, D.J., Boudousquie, A., Copeland, K., Young, V., Kalinowski, S., & Vaughn, S. (2006). Effects of accommodations on high-stakes testing for students with reading disabilities. *Exceptional Children, 72* (2), 136-150.
- Fuchs, L. S., Fuchs, D., Eaton, S. B., Hamlett, C. L., & Karns, K. M. (2000). Supplementing teacher judgments of mathematics test accommodations with objective data sources. *School Psychology Review, 29*(1), 65-85.
- Helwig, R., Rozek-Tedesco, M.A., & Tindal, G. (2002). An oral versus a standard administration of a large-scale mathematics test. *The Journal of Special Education, 36* (1), 39-47.
- Huynh, H., & Barton, K.E. (2006). Performance of students with disabilities under regular and oral administrations of a high-stakes reading examination. *Applied Measurement in Education, 19* (1), 21-39.
- Huynh, H., Meyer, J.P., & Gallant, D.J. (2004). Comparability of student performance between regular and oral administrations for a high-stakes mathematics test. *Applied Measurement in Education, 17* (1), 39-57.
- Janson, I.B. (2002). The effects of testing accommodations on students' standardized test scores in a northeast Tennessee school system (Doctoral dissertation, East Tennessee State University, 2002). *Dissertation Abstracts International, 63/02*. 557.
- Johnstone, C.J., Altman, J., Thurlow, M.L., & Thompson, S.J. (2006). *A summary of research on the effects of test accommodations: 2002 through 2004* (Technical Report 45). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Lazarus, S.S., Thurlow, M.L., Lail, K.E., Eisenbraun, K.D., & Kato, K. (2006). *2005 state policies on assessment participation and accommodations for students with disabilities* (Synthesis Report 64). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
-

- Lesaux, N. K., Pearson, M. R., & Siegel, L. S. (2006). The effects of timed and untimed testing conditions on the reading comprehension performance of adults with reading disabilities. *Reading and Writing, 19*, 21-48.
- Koretz, D., & Hamilton, L. (2000). Assessment of students with disabilities in Kentucky: Inclusion, student performance, and validity. *Educational Evaluation and Policy Analysis, 22* (3), 255-272.
- Mandinach, E. B., Bridgeman, B., Cahalan-Laitusis, C., & Trapani, C. (2005). The impact of extended time on SAT test performance. *Research Report No 2005-8*. New York, NY: The College Board.
- McKevitt, B.C., & Elliott, S.N. (2003). Effects and perceived consequences of using read aloud and teacher-recommended testing accommodations on a reading achievement test. *The School Psychology Review, 32* (4) 583-600.
- Meloy, L.L., Deville, C., & Frisbie, D. (2002). The effect of a read aloud accommodation on test scores of students with and without a learning disability in reading. *Remedial and Special Education, 23* (4), 248-255.
- National Accessible Reading Assessment Projects (2006). *Defining reading proficiency for accessible large-scale assessments: Some guiding principles and issues*. Minneapolis, MN: University of Minnesota, National Accessible Reading Assessments. Available at www.narap.info.
- Schnirman, R.K. (2005). The effect of audiocassette presentation of the performance of students with and without learning disabilities on a group standardized math test. *Dissertation Abstracts International, 66* (6), 2172 A.
- Schulte, A.G., Elliott, S.N., & Kratochwill, R.R. (2001). Effects of testing accommodations on standardized mathematics test scores: An experimental analysis of the performances of students with and without disabilities. *School Psychology Review, 30* (4), 527-547.
- Sireci, S., Scarpati, S., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. *Review of Educational Research, 75*, 457-490.
- Thompson, S., Blount, A., & Thurlow, M. (2002). *A summary of research on the effects of test accommodations: 1999 through 2001* (Technical Report 34). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thompson, S.J., Morse, A.B., & Thurlow, M.L. (2005). *Access assistants for state assessments: A study of state guidelines for scribes, readers, and sign language interpreters* (Synthesis Report 58). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Thornton, A.E., Reese, L.M., Pashley, P.J., & Dalessandro, S.P. (2002). *Predictive validity of accommodated LSAT scores*. Pennsylvania: Law School Admission Council.
- Thurlow, M.L. (2007). State policies and accommodations: Issues and implications. In C. Cahalan-Laitusis & L. Cook (eds.), *Accommodating students with disabilities on state assessments: What works?* Arlington, VA: Council for Exceptional Children.
- Thurlow, M.L., McGrew, K.S., Tindal, G., Thompson, S.L., Ysseldyke, J.E., & Elliott, J.L. (2000). *Assessment accommodations research: Considerations for design and analysis* (Technical Report 26). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes. Available from www.nceo.info/OnlinePubs/Technical26.htm
- Thurlow, M.L., Moen, R., & Altman, J. (2006). Annual performance reports: 2003-2004 state assessment data (available on the NCEO Web site at <http://www.nceo.info/OnlinePubs/APRsummary2006.pdf>). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
-

- Thurlow, M.L., Scott, D.L., & Ysseldyke, J.E. (1995). *A compilation of states' guidelines for accommodations in assessments for students with disabilities* (Synthesis Report 18). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Tindal, G. (1998). Models for understanding task comparability in accommodated testing. A publication for the Council of Chief State School Officers, Washington, DC. Retrieved May 19, 2006, from the World Wide Web: <http://education.umn.edu/nceo/OnlinePubs/Accomm/TaskComparability.htm>
- Tindal, G. (2002). *Accommodating mathematics testing using a videotaped, read-aloud administration*. Washington, DC: Council of Chief State School Officers.
- Tindal, G., & Fuchs, L. (1999). *A summary of research on test changes: An empirical basis for defining accommodations*. Lexington, KY: University of Kentucky, Mid-South Regional Resource Center.
- U.S. Department of Education. (2004, April 28). *Standards and assessment peer review guidance: Information and examples for meeting the requirements of the No Child Left Behind Act of 2001*. Washington, DC: Office of Elementary and Secondary Education.
- Weston, T.J. (2003). *NAEP validity studies: The validity of oral accommodation testing*. Washington, DC: National Center for Education Statistics.
- Woods, K. (2004). Deciding to provide a reader in examinations for the general certificate of secondary education (GCSE): Questions about validity and inclusion. *British Journal of Special Education*, 31 (3), 122-124.
- Zenisky, A. L., & Sireci, S. G. (in press). *A summary of the research of the effects of test accommodations: 2005 through 2006* (NCEO Technical Report). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Acknowledgment

The preparation of this paper was supported, in part, by a cooperative agreement between the U.S. Department of Education, Research to Practice Division of the Office of Special Education Programs (#H326G050007) and the University of Minnesota. Opinions expressed in this paper do not necessarily reflect those of the U.S. Department of Education or Offices within it.
